Neurocomputing 275 (2018) 403–417

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Kernelized random KISS metric learning for person re-identification

Cairong Zhao [a,b,c,1,*], Yipeng Chen [a,1], Xuekuan Wang [a], Wai Keung Wong [c,d], Duoqian Miao [a,b], Jingsheng Lei [e]

[a] Department of Computer Science and Technology, Tongji University, Shanghai, China
[b] Key Laboratory of Embedded System and Service Computing (Tongji University), Ministry of Education, Shanghai 201804, China
[c] Institute of Textiles and Clothing, The Hong Kong Polytechnic University, Hong Kong, China
[d] The Hong Kong Polytechnic University, Shenzhen Research Institute, China
[e] College of computer science and technology, Shanghai University of Electric Power, China

ARTICLE INFO

Article history:
Received 17 February 2017
Revised 29 July 2017
Accepted 25 August 2017
Available online 8 September 2017

Communicated by Dr Jiwen Lu

Keywords:
Metric learning
Kernelized feature
Random subspace ensemble

ABSTRACT

Person re-identification is critical for human tracking in the video surveillance which has attracted more and more attention in recent years. Various recent approaches have made great progress in re-identification performance using metric learning techniques and among them, Keep It Simple and Straightforward (KISS) metric learning method has shown remarkably importance because of its simpleness and high-efficiency. The KISS method is based on an assumption that the differences between feature pairs obey the Gaussian distribution. However, for most existing features of person re-identification, the distributions of differences between feature pairs are irregular and undulant. Therefore, prior to the Guassian based metric learning step, it's important to augment the Guassian distribution of data without losing discernment. Moreover, most metric learning methods were greatly influenced by the small sample size (SSS) problem and the KISS method is no exception, which causing the inexistence of inverse of covariance matrices. To solve the above two problems, we present Kernelized Random KISS (KRKISS) metric learning method. By transforming the original features into kernelized features, the differences between feature pairs can better fit the Gaussian distribution and thus they can be more suitable for the Guassian assumption based models. To solve the inverse of covariance matrix estimation problem, we apply a random subspace ensemble method to obtain exact estimation of covariance matrix by randomly selecting and combining several different subspaces. In each subspace, the influence of SSS problem can be minimized. Experimental results on three challenging person re-identification datasets demonstrate that the KRKISS method significantly improves the KISS method and achieves better performance than most existing metric learning approaches.

1. Introduction

Person re-identification is a sub-problem in automated video surveillance which has attracted more and more attention in the past five years because of its significance [1,2]. Given an image of one person, the objective of person re-identification is to match him/her among a large number of images of pedestrians which are captured by some different cameras. As humans, we do it all the time without paying much effort, but it's an extremely difficult task for computer because there are many factors that may influence the performance. The images of pedestrians are always low-quality, which causes the traditional biometrics, such as face [24,38], gait [39] and iris [40], to be not available. Moreover, the appearances of pedestrians often change across camera views due to wide variations in viewpoints, illumination, and pose. The occlusions in the images also bring more difficulty to it. To deal with this problem, researchers have proposed many different methods in which the most common research direction is feature representation. It aims to extract robust and effective features from each pedestrian image. The features can be hand-crafted [4–6,18-20,29] or learned [3,17,28,30–34,44,59,60] based on visual information like color, texture, and shape. It has been proved that among all the visual information, color plays the most important role for person re-identification.

After extracting distinguishing features, the next step is naturally to compute the distances between different feature pairs and choose the smallest one as the matching image pair. Different distances, such as Euclidean distance, cosine distance, and Mahalanobis distance, are usually used for pedestrian matching. The

* Corresponding author at: Department of Computer Science and Technology, Tongji University, Shanghai, China.
E-mail address: zhaocairong@tongji.edu.cn (C. Zhao).
[1] The authors contribute equally to this work.

http://dx.doi.org/10.1016/j.neucom.2017.08.064
0925-2312/

procedure of learning a discriminant distance function is called metric learning. Researchers have proposed many metric learning algorithms for person re-identification which can be summarized to several groups according to their basic theories. For instance, based on the theory of Fisher Discriminant Analysis (FDA), some metric learning methods such as, Local Fisher Discriminant Analysis (LFDA) [7], Marginal Fisher Analysis (MFA) [8] and Cross-view Quadratic Discriminant Analysis (XQDA) [6], were proposed. Most of them tried to project descriptors into a more discriminatory subspace based on an eigenvalue resolution. By adding special constraints to the distance functions, some methods like Pairwise Constrained Component Analysis (PCCA) [9], Large Margin Nearest Neighbor (LMNN) [10] and Locally-Adaptive Decision Functions (LADF) [11] were proposed. Different constraints in the methods show different performances when they are applied for person re-identification. Some kernel-based metric learning methods also have been proposed to improve the identification ability of features because of the non-linearity in the appearance of pedestrians [5,8,21,35,41–43]. The benefits of kernel-based methods for person re-identification are twofold: (1) True matches and wrong matches become more separablein the nonlinear space; (2) the kernel can be chosen flexibly. In addition, value difference metric (VDM) [49] and its varieties [50-58] have been proposed to find reasonable distance between each pair of instances with nominal attributes only.

It is specially worth mentioning that based on the maximum likelihood (ML) estimation, Keep It Simple and Straightforward (KISS) metric learning method [12] was proposed. The KISS metric learning method is both efficient and effective by considering a log likelihood ratio test of two Gaussian distributions as the distance between a feature pair. With a simplified and efficient solution, the covariance matrix of Mahalanobis distance is obtained by computing the difference between the inverse of two covariance matrices. Although the KISS metric learning method has been widely applied to the person re-identification because of its simpleness and effectivenes36s, there are still two problems. Firstly, there is an assumption of the KISS metric learning method that pairwise differences are agreeing with Gaussian distribution, if not, the performance would decrease. Second, most metric learning methods suffer from SSS (small sample size) problem because the size of samples is always much smaller than the dimensionality of features, and the KISS metric learning method is no exception. In the training procedure of the KISS method, the covariance matrices may be singular when the training size of data is small, resulting in the inexistence of the inverse of covariance matrices. How to augment the Guassian distribution of data without losing discernment and how to accurately estimate the inverse of covariance matrix in the case of small samples are still open problems which exsit in various fields. To overcome the inverse of covariance matrix estimation problem, unsupervised dimensionality reduction is required as the most common method. In [12], before applying the KISS metric learning method, the PCA (Principal Component Analysis) was used to project the concatenated descriptors into a 34-dimensional subspace which was greatly less than the original feature dimension. However, in the procedure of dimension reduction, the abandon dimensionalities also have the ability of distinguishing pedestrians, and thus the unsupervised dimensionality reduction will make the learned distance less discriminative. There are some improved metric learning methods based on the KISS method such as Regularized Smoothing KISS (RS-KISS) [13], minimum classification error-KISS (MCE-KISS) [14], Regularized Bayesian Metric Learning (RBML) [15] and Dual-Regularized KISS (DR-KISS) [36]. Although improvement of performance can be seen, the problems mentioned above are still unsolved.

In this paper, based on the KISS method we propose a novel metric learning method called Kernelized Random KISS (KRKISS) in which a special kernel technique is used to augment the Gaussian distributionof data by transforming the original features into kernelized features, and a random subspace ensemble method is applied tosolve the inverse of covariance matrix estimation problem in the case of small samples. It contains the advantages of the KISS metric learning method and the kernel theory, and avoids using dimensionality reduction techniques or regularization techniques which will cause a loss to discernment of features. We have introduced above that the KISS method was proposed under the assumption of Gaussian distribution. When the original features are transformed into kernelized features by our proposed method, they can better fit Gaussian distribution and obtain more discrimination for person re-identification. It is the inspiration and motivation that why we utilize the transformation to improve the KISS metric learning method. As far as we know, it is the first of its kind to improves the KISS method by augmenting the Guassian model of data. Moreover, the covariance matrix usually be singular when the size of samples is small, which would cause the inexistence of the inverse of covariance matrix. The random subspace ensemble method we applied to obtain exact estimation of covariance matrix is simple but very efficient, which decreases the feature dimension by randomly selecting and combining several different subspaces. In each subspace, the influence of SSS problem can be minimized. It also works to other covariance matrix based metric learning methods. Experiments with five kinds of features (LOMO [6], kCCA [5], SCNCD [4], ELF18 [19] and FFN [17]) on different datasets show that our proposed KRKISS metric learning algorithm can improve the identification accuracy obviously and get better performance than other state-of-the-art metric learning methods in most of experiments.

The main contributions of our work are summarized as below:

(1) We augment the Gaussian distribution of data by transforming the original feature into special kernelized feature. By doing so, the differences between feature pairs can better fit the Gaussian distribution, which is more aligned with the metric learning methods based on Gaussian distribution assumption.
(2) We apply a random subspace ensemble method to take the place of regularization methods in the procedure of estimating the inverse of covariance matrix. In each randomly selected subspace, the influence of SSS problem can be minimized and the inverse of covariance matrix can be estimated more accurately.

The rest of the paper is organized as follows. Section 2 introduces some related works about person re-identification. Section 3 proposes our Kernelized Random KISS (KRKISS) metric learning method. Experiments on three public datasets are described in Section 4. Finally, the conclusion is draw in Section 5.

2. Related work

The proposed methods for person re-identification can be generally categorized into two groups which have been briefly described in Section 1. The first group of methods focuses on designing features which are robust to viewpoint change, illumination change and other variations in pedestrian's images. There are many useful features like the ensemble of localized features (ELF) [3], salient color names based color descriptor (SCNCD) [4], Local Maximal Occurrence (LOMO) [6] and hierarchal Gaussian descriptor (GOG) [29] which have been proposed and applied in person re-identification successfully. However, they are still not robust enough to huge variations in pedestrian's images. Therefore, more and more researchers turn attention to the second group of methods (metric learning) which focuses on learning a robust distance or similarity function to deal with the matching problem.

Many effective metric learning algorithms have been proposed and widely applied to solve the person re-identification problem. Davis et al. [37] proposed information-theoretic metric learning (ITML) from the perspective of information theoretic. Weinberger et al. [10] proposed the large-margin nearest neighbor metric (LMNN) to improve the performance of the kNN classification. In [9], an algorithm for learning distance metric called Pairwise Constrained Component Analysis (PCCA) was proposed. It projected the features into a low-dimensional space where the distance between pairs of data points respects the desired constraints. In addition, some kernel based metric learning methods have been proposed in recent years. Xiong et al. [8] proposed kernel LFDA metric learning method which performs better than LFDA [7]. Lisanti et al. [5] applied a learning technique based on Kernel Canonical Correlation Analysis (KCCA) which finds a common subspace between the proposed descriptors extracted from disjoint cameras and then projects them into a new descriptors space. Hu et al. [48] proposed a new deep transfer metric learning (DTML) method to learn a set of hierarchical nonlinear transformations by transferring discriminative knowledge from the labeled source domain to the unlabeled target domain. Duan et al. [47] proposed a deep localized metric learning (DLML) approach by learning multiple fine-grained deep localized metrics. In [21], kernel technique was utilized to learning a kernelized discriminative null space.

Moreover, Köstinger et al. [12] proposed the KISS metric learning algorithm which considered a log likelihood ratio test of two Gaussian distributions and obtained a simplified and efficient solution to solve it. There are also some improved methods based on the KISS method. In [13], a regularized smoothing KISS metric learning (RS-KISS) was presented which seamlessly integrated smoothing and regularization techniques to estimate covariance matrices. In [14], the KISS metric learning method was improved based on the minimum classification error (MCE). Tao et al. [36] proposed a Dual-Regularized KISS (DR-KISS) metric learning by regularizing the two covariance matrices and theoretically demonstrated that the proposed regularization method is robust for generalization. However, the above KISS based improved methods focus on solving the SSS problem. They may still perform poorly when the features do not obey the Gaussian Distribution.

In this paper, we focus on the second group of methods, i.e. metric learning. Aiming at augmenting Guassian distribution of data and exactly estimating the inverse of covariance matrix, we propose Kernelized Random KISS (KRKISS) metric learning method.

3. Proposed approach

3.1. Kernelizd random KISS metric learning (KRKISS)

The person re-identification problem can be formulated as follows. Let $(x_i, y_i)_{i=1}^m$ denote all training samples, where x_i is the i-th sample, y_i is the identity of the corresponding person and m is the number of training samples. The goal of metric learning methods is to learn a generic distance function $d(\cdot)$ from the training samples that outputs smaller value for samples of the same person and larger value for different people. During the test phase, given a probe pedestrian image and a set of gallery images, they are transformed into descriptors firstly. Then we use the distance function $d(\cdot)$ to measure distance from all of them, and rank the gallery images according to their distance to the probe image. Usually we use the Mahalanobis distance function to measure the distance between two samples x_i and x_j, such as

$$d(x_i, x_j) = (x_i - x_j)^T M (x_i - x_j)$$

thus we need to learn a robust and discriminative matrix M from training samples. In the KISS metric learning method [12], under a assumption that the differences between features pairs obey

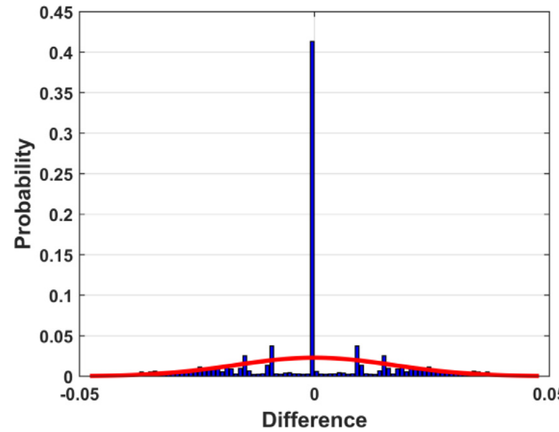


Fig. 1. The typical distribution of differences between feature pairs. The blue bar represents the actual distribution and the red line represents the computed Gaussian distribution. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Guassian distribution, the matrix M was calculated by

$$M = \Sigma_1^{-1} - \Sigma_0^{-1}$$

where Σ_1 is the covariance matrix of differences between same person and Σ_0 is the covariance matrix of differences between different people.

It is an simple and efficient metric learning algorithm, however, the commonly used features for person re-identification do not obet the assumption that the differences between feature pair are Gaussian distributed. For example, the typical distribution of differences between feature pairs are shown in Fig. 1. As we can see, the actual distribution is at odds with the computed Gaussian distribution, which causes a significant decrease in performance of Guassian distribution based metric learning methods. The Guassian distribution assumption puts forward high demand to the feature representation. Moreover, the sample size is small in most cases, and thus the estimation of inverse of covariance matrix is inaccurate, or not exist. Some regularization methods have been proposed to solve it [13,14,15,36]. However, the existing regularization methods are still not robust enough because the weights of regularization are hand-crafted. To solve those two problems, we propose Kernelized Random KISS (KRKISS) metric learning method. For convenience, Table 1 lists frequently used notations and descriptions in the paper.

The first procedure of KRKISS is to transform the original features into kernelized features. Given a function $\Phi(x_i)$ which maps the sample x_i into an unknown space. Different with the common usage of kernel trick, we utilize the inner products of $\Phi(x_i)$ and a training sample feature $\Phi(x_j)$ as a novel feature of sample x_i. For example, $\Phi(x_i)^T\Phi(x_j)$ denotes the j-th novel feature, and thus the sample x_i can be transformed into $[\Phi(x_i)^T\Phi(x_1), \Phi(x_i)^T\Phi(x_2), \ldots, \Phi(x_i)^T\Phi(x_m)]$ when there are m training samples. Given a kernel function $k(x_i, x_j) = \langle\Phi(x_i), \Phi(x_j)\rangle = \Phi(x_i)^T\Phi(x_j)$, x_i can be transformed into kernelized features $\tilde{x}_i = \mathbf{K}(x_i, \mathbf{X}) = [k(x_i, x_1), k(x_i, x_2), \ldots, k(x_i, x_m)]^T \in R^{m\times 1}$. In the same way, the training samples $\mathbf{X}$ and the testing samples $\mathbf{Z}$ can be transformed into $\tilde{\mathbf{X}} = [\mathbf{K}(x_1, \mathbf{X}), \mathbf{K}(x_2, \mathbf{X}), \ldots, \mathbf{K}(x_m, \mathbf{X})] \in \mathbf{R}^{m\times m}$ and $\tilde{\mathbf{Z}} = [\mathbf{K}(z_1, \mathbf{X}), \mathbf{K}(z_2, \mathbf{X}), \ldots, \mathbf{K}(z_n, \mathbf{X})] \in \mathbf{R}^{m\times n}$ respectively. Unless stated, otherwise, RBF kernel is used in our method with kernel width determined automatically using the mean pairwise distance of samples. After transforming the original features into kernelized features, the actual distribution of differences between kernelized feature pairs better fit the computed Gaussian distribution.

Table 1
Frequently used notations and descriptions in the paper.

Notation	Description	Notation	Description
$\mathbf{X}/x_i$	Training samples	$\mathbf{Z}/z_i$	Testing samples
m	Number of training samples	n	Number of testing samples
$\tilde{\mathbf{X}}/\tilde{x}_i$	Kernelized features of training samples	$\tilde{\mathbf{Z}}/\tilde{z}_i$	Kernelized features of testing samples
$\tilde{x}_{ij}$	Difference of kernelized feature pair and $\tilde{x}_j$	$\tilde{z}_{ij}$	Difference of kernelized feature pair $\tilde{z}_i$ and $\tilde{z}_j$

Moreover, there are many advantages of kernelized features which will be described in detail and analyzed in next section.

Given a pair of testing samples z_i and z_j that we need to decide if they are the same person. They can be transformed to kernelized features $\tilde{z}_i$ and $\tilde{z}_j$ according to the above method. Let H_0 denotes the hypothesis that the testing feature pair ($\tilde{z}_i$ and $\tilde{z}_j$) is dissimilar ($\tilde{z}_i$ and $\tilde{z}_j$ are different people), H_1 denotes the hypothesis that the feature pair ($\tilde{z}_i$ and $\tilde{z}_j$) is similar ($\tilde{z}_i$ and $\tilde{z}_j$ are the same people). If $\tilde{z}_{ij}$ satisfies the Gaussian distribution, we have

$$p(\tilde{z}_{ij}|H_k) = \frac{1}{(2\pi)^{m/2}|\Sigma_k|^{1/2}} \exp\left(-\frac{1}{2}\tilde{z}_{ij}^T\Sigma_k^{-1}\tilde{z}_{ij}\right) \tag{1}$$

where $k \in \{0, 1\}$, $p(\tilde{z}_{ij}|H_0)$ and $p(\tilde{z}_{ij}|H_1)$ denotes the probability of $\tilde{z}_{ij}$ under the hypothesis of H_0 and H_1 respectively, Σ_0 is the covariance matrix of interpersonal differences and Σ_1 is the covariance matrix of intrapersonal differences.

The two covariance matrices can be estimated by the training samples. Define y_{ij} as the indicated variable of two training samples $\tilde{x}_i$ and $\tilde{x}_j$: $y_{ij} = 1$ if $\tilde{x}_i$ and $\tilde{x}_j$ are the same person and $y_{ij} = 0$, otherwise. So we have

$$\Sigma_0 = \frac{1}{N_0}\sum_{y_{ij}=0}\tilde{x}_{ij}\tilde{x}_{ij}^T = \frac{1}{N_0}\sum_{y_{ij}=0}(\tilde{x}_i - \tilde{x}_j)(\tilde{x}_i - \tilde{x}_j)^T$$

$$\Sigma_1 = \frac{1}{N_1}\sum_{y_{ij}=1}\tilde{x}_{ij}\tilde{x}_{ij}^T = \frac{1}{N_1}\sum_{y_{ij}=1}(\tilde{x}_i - \tilde{x}_j)(\tilde{x}_i - \tilde{x}_j)^T \tag{2}$$

where N_0 denotes the number of interpersonal differences, N_1 denotes the number of intrapersonal differences.

In order to decide if $\tilde{z}_i$ and $\tilde{z}_j$ are the same person, the logarithm of the ratio between two posterior probabilities is defined as follows:

$$\delta(\tilde{z}_{ij}) = \log\left(\frac{p(H_0|\tilde{z}_{ij})}{p(H_1|\tilde{z}_{ij})}\right) \tag{3}$$

According to the Bayes formula, (3) can be transformed to

$$\delta(\tilde{z}_{ij}) = \log\left(\frac{p(\tilde{z}_{ij}|H_0)}{p(\tilde{z}_{ij}|H_1)}\right) + \log\left(\frac{p(H_0)}{p(H_1)}\right) \tag{4}$$

Given (1), (4) can be rewritten as

$$\delta(\tilde{z}_{ij}) = \frac{1}{2}\tilde{z}_{ij}^T(\Sigma_1^{-1} - \Sigma_0^{-1})\tilde{z}_{ij} + \frac{1}{2}\log\left(\frac{|\Sigma_1|}{|\Sigma_0|}\right) + \log\left(\frac{p(H_0)}{p(H_1)}\right) \tag{5}$$

So $\tilde{z}_i$ and $\tilde{z}_j$ are the same people while the value of $\delta(\tilde{z}_{ij})$ is negative and a positive value of $\delta(\tilde{z}_{ij})$ indicates $\tilde{z}_i$ and $\tilde{z}_j$ are different people. By dropping the constant in (5), we have

$$\delta(\tilde{z}_{ij}) = \tilde{z}_{ij}^T(\Sigma_1^{-1} - \Sigma_0^{-1})\tilde{z}_{ij} \tag{6}$$

The smaller the value of $\delta(\tilde{z}_{ij})$ is, the bigger the probability of $\tilde{z}_i$ and $\tilde{z}_j$ belong to same person. It is similar with the Mahalanobis distance in form, and thus the distance between two testing samples z_i and z_j is

$$d(z_i, z_j) = (\tilde{z}_i - \tilde{z}_j)^T(\Sigma_1^{-1} - \Sigma_0^{-1})(\tilde{z}_i - \tilde{z}_j) \tag{7}$$

The dimension of kernelized feature is the same as the number of training samples, which is greatly less than the dimension of original feature in most of scenes. However, the SSS problem still remains, resulting in the inexistence of the inverse matrices Σ_0^{-1} and Σ_1^{-1}. The common solution is to add a regularization process for covariance matrix such as

$$\Sigma_0 = (1-\alpha_0)\Sigma_0 + \alpha_0\frac{\text{trace}(\Sigma_0)}{d}I$$

$$\Sigma_1 = (1-\alpha_1)\Sigma_1 + \alpha_1\frac{\text{trace}(\Sigma_1)}{d}I \tag{8}$$

where α_0 and α_1 denotes the weights of regularization. The significance of this regularization method has been demonstrated theoretically in [36]. However it is not robust enough because the parameters α_0 and α_1 are hand-controlled.

To address this problem, we propose a random subspace ensemble method which projects the training features into different randomly selected subspaces. Then the inverse matrices are computed by all the traning samples in each subspace. When the dimension of subspace is greatly less than the number of training samples, we can obtain the inverse matrices without the SSS problem. The final distance is computed by

$$d(z_i, z_j) = \sum_{k=1}^{L}(D_k(\tilde{z}_i) - D_k(\tilde{z}_j))^T M_k(D_k(\tilde{z}_i) - D_k(\tilde{z}_j)) \tag{9}$$

$$\begin{aligned} M_k &= \Sigma_1^{-1} - \Sigma_0^{-1} \\ &= \frac{1}{N_1}\sum_{y_{ij}=1}D_k(\tilde{x}_{ij})D_k(\tilde{x}_{ij})^T - \frac{1}{N_0}\sum_{y_{ij}=0}D_k(\tilde{x}_{ij})D_k(\tilde{x}_{ij})^T \end{aligned} \tag{10}$$

where L denotes the number of selected subspaces and D_k denotes a randomly selected subspace. The diagram of KRKISS metric learning method is shown in Fig 2 and the complete KRKISS algorithm is summarized in Algorithm 1.

3.2. Discussion of kernelized random KISS method

The Kernelized Random KISS (KRKISS) method has the main advantages as below:

(1) When the original features are transformed into kernelized features, the differences between feature pairs can better fit Gaussian distribution.

For example, given the LOMO feature $[d_1, d_2, \ldots, d_{26960}]^T = X \in R^{26960\times632}$ extracted from VIPeR dataset [25], where 26,960 is the dimensionality of the LOMO feature, 632 is the number of samples and $d_j \in R^{1\times632}$ denotes the j-th dimension data. For every one-dimensional data d_k, it is able to calculate two sets of differences dif_0 and dif_1 according to the labels of samples, where dif_0 denotes a set of differences between interperson feature pairs and dif_1 denotes a set of differences between intraperson feature pairs. Since the means of dif_0 and dif_1 are both zero, we can obtain two Gaussian distributions by calculating the variances of dif_0 and dif_1 respectively. Comparing the concrete probability distributions of differences with the computed Gaussian distribution, we can consider the set of differences obeys the Gaussian distribution if they are similar in shape. The typical probability distributions of four original features (LOMO [6], kCCA [5], SCNCD [4] and ELF18 [19]) are shown in Fig 3a and b. The blue bars are concrete probability distributions of differences and the red line is the computed

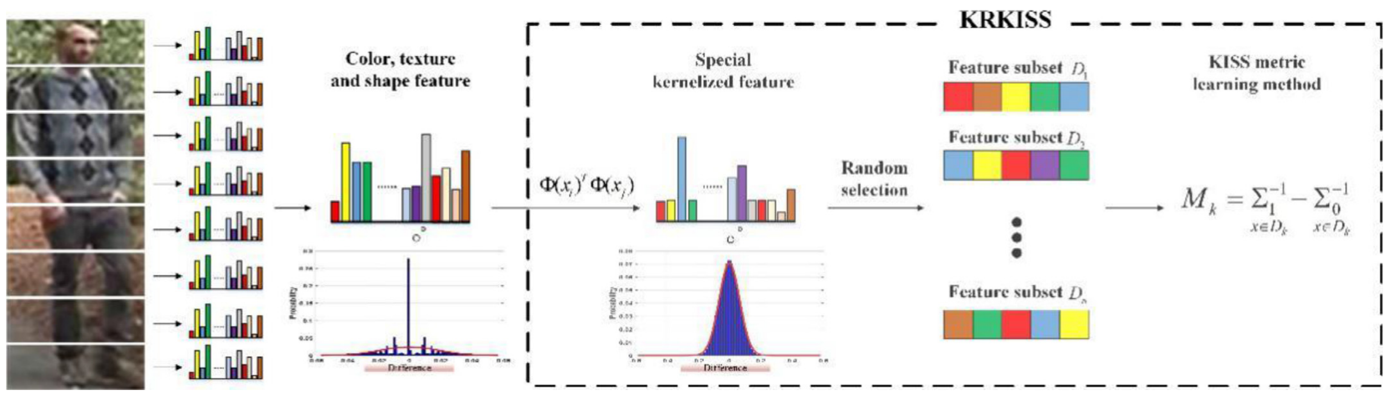


Fig. 2. The diagram of KRKISS metric learning method.

Algorithm 1
Kernelized Random KISS (KRKISS).

Input: Training samples $(x_i, y_i)_{i=1}^{m}$, testing samples $(z_i)_{i=1}^{n}$, the number of se lected subspaces L and the dimension of subspace d
Output: Some Mahalanobis matrices M_k, $(k = 1, 2, \ldots, L)$ and the distance be tween two test samples z_i and z_j
while $i < m$
$i \leftarrow i + 1$;
Calculate the kernelized feature of ith training sample by $\tilde{x}_i = \mathbf{K}(x_i, \mathbf{X}) = [k(x_i, x_1), k(x_i, x_2), \ldots, k(x_i, x_m)]^T$;
end while
while $t < n$
$t \leftarrow t + 1$;
Calculate the kernelized feature of t-th testing sample by $\tilde{z}_t = \mathbf{K}(z_t, \mathbf{X}) = [k(z_t, x_1), k(z_t, x_2), \ldots, k(z_t, x_m)]^T$;
end while
while $k < L$
Randomly select a d-dimension subspace D_k from kernealized feature space;
Calculate the k-th Mahalanobis matrix M_k according to Eq. (10);
end while
Calculate the distance between two test samples z_i and z_j according to Eq. (9)

Table 2
The Lilliefors test results of two features.

Features /Lilliefors test	LOMO		kCCA	
	Original features	Kernelized features	Original features	Kernelized features
Lilliefors test of dif_1	0.0462	0.3497	0.4673	0.6139
Lilliefors test of dif_0	0.0355	0.3956	0.4691	0.8149

Table 3
The variance ratio of two features.

Features	LOMO		kCCA	
	Original features	Kernelized features	Original features	Kernelized features
variance ratio $\frac{\sigma_0}{\sigma_1}$	1.1502	1.4389	1.2429	2.4177

Gaussian distribution. As we can see in Fig 3a and b, the probability distributions of the differences between original feature pairs are irregular and most of them are concentrated around the zero value. The computed Gaussian distributions do not fit the probability distributions of the differences since the values of variance are higher when the distributions are irregular. In the same way, it is able to calculate the differences between kernelized feature pairs and the typical probability distributions of four kernelized features (LOMO, kCCA, SCNCD, ELF18), as shown in Fig 3c and d. When the original features are transformed into kernelized feature, the probability distributions of the differences will be more regular and better fit to the computed Gaussian distribution in shape.

Moreover, we have applied the Lilliefors test [45] to test the null hypothesis that the data come from a normally distributed population. The Lilliefors test is a normality test based on the Kolmogorov-Smirnov test with mean and variance unknown. We set the value of $\frac{n_0}{N}$ to represent the degree of distribution, where n_0 denotes the number of one-dimensional feature which is consistent with the outcome of Lilliefors test, N denotes the dimensionality of features. The value of $\frac{n_0}{N}$ is larger and the features better fit to Gaussian distribution. The results are shown in Table 2. As we can see in Table 2, the value of $\frac{n_0}{N}$ increases significantly when the original features are transformed into kernelized feature. It demonstrates that the kernelized feature better fit to the Gaussian distribution and is more suitable for the Guassian distribution assumption based metric methods.

(2). The kernelized features always have higher recognition capability in the non-linear space.

The variance ratio of two differences sets dif_0 and dif_1 can be seen as a performance index of discernment [6]. The distributions of dif_0 and dif_1 are shown in Fig 4. σ_0 denotes the variance of dif_0 and σ_1 denotes the variance of dif_1. The higher the variance ratio $\frac{\sigma_0}{\sigma_1}$ is, the more discriminative the feature is. The value of variance ratio for every one-dimensional feature was calculated and the average value was taken. The results are shown in Table 3. Since the mean variance ratio of kernelized feature is higher than the original feature, we consider that when the original features are transformed into the kernelized features, they would become more distinguishable.

(3) The robustness to the variation of parameters.

There are two parameters in KRKISS method, the number of selected subspaces L and the dimension of subspace d. Experientially we set $d = \frac{m}{4}$ to guarantee the existence of the inverse of covariance matrices in subspace. Next, we conduct some experiments with different parameters to compare the influence of L in KRKISS method. Fig 5 shows the experimental results of LOMO feature and kCCA feature respectively. The blue line represents the impact

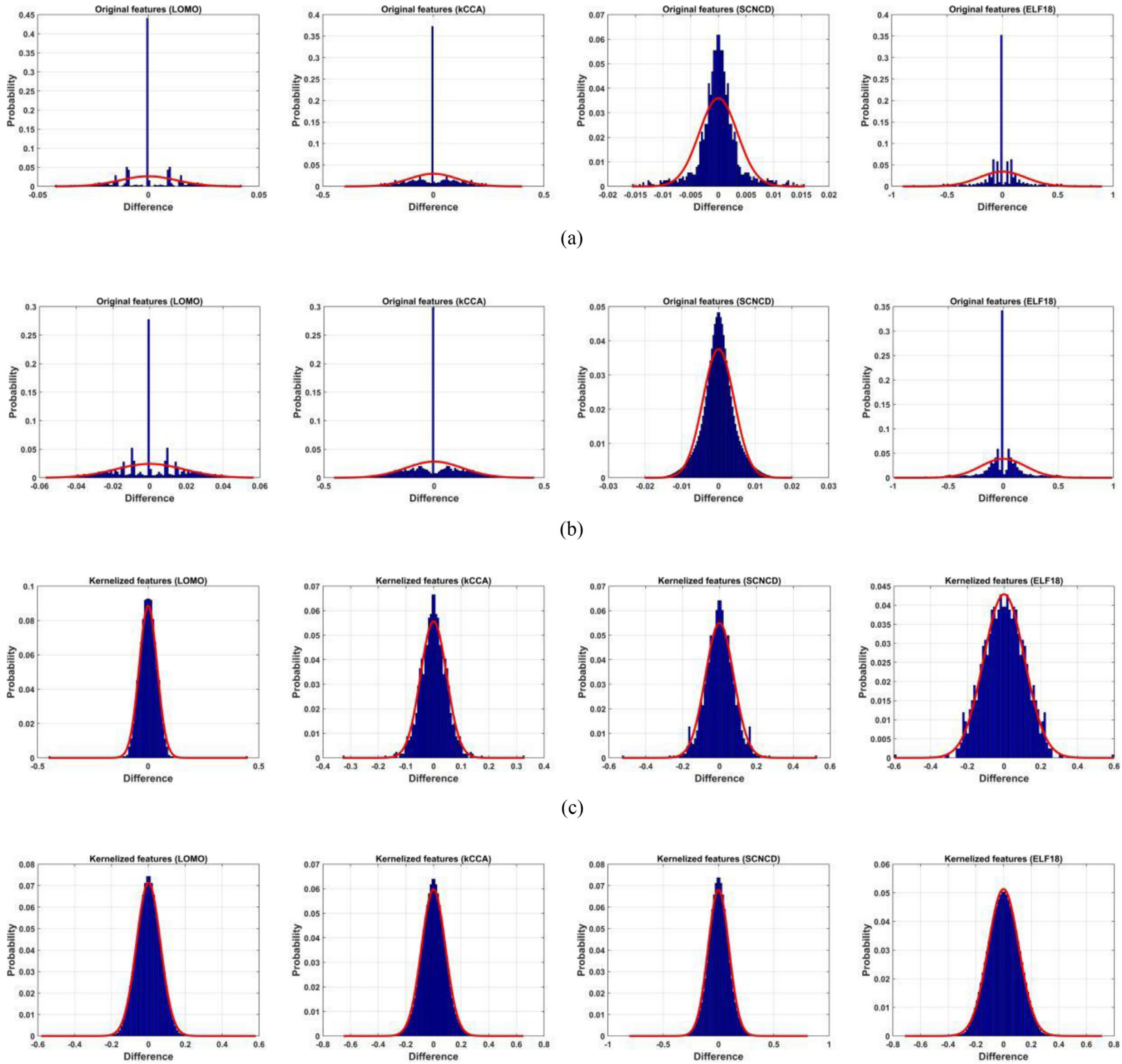


Fig. 3. The probability distributions. The blue bars are concrete probability distributions of differences and the red line is the computed Gaussian distribution of differences. (a) Differences of original features between intraperson image pairs. (b) Differences of original features between interperson image pairs. (c) Differences of kernelized features between intraperson image pairs. (d) Differences of kernelized features between interperson image pairs. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of subspaces number on rank-1 matching rate and the red line represents the impact of regularized weight on rank-1 matching rate.

As we can see in Fig 5, the matching rate of regularized method has obvious fluctuations which cause the difficulty to choose a better parameter. Instead, the matching rate of random subspace ensemble method always achieves near the maximum when the number of subspaces increases,. Taking into account of the efficiency of algorithm, we set $L = \frac{d}{2}$, and thus all the parameters in the KRKISS method need not to be adjusted manually.

(4) Low computational complexity.

Suppose the m training samples consist of a images captured from one view and b images captured from the other view ($m = a + b$). In KRKISS method, the computation of feature transform requires $O(dm^2)$ multiplication operations and for each randomly selected subspace, the computation of the two covariance matrices Σ_I and Σ_E require $O(Nkm^2)$ and $O(abm^2)$ multiplication operations, respectively, where $N = \max(a, b)$, d denotes the dimension of original feature and k represents the average number of images in each person. Therefore, the total computation of the two covariance matrices Σ_I and Σ_E in KRKISS method require $O(LNkm^2)$ and $O(Labm^2)$ multiplication operations, respectively.

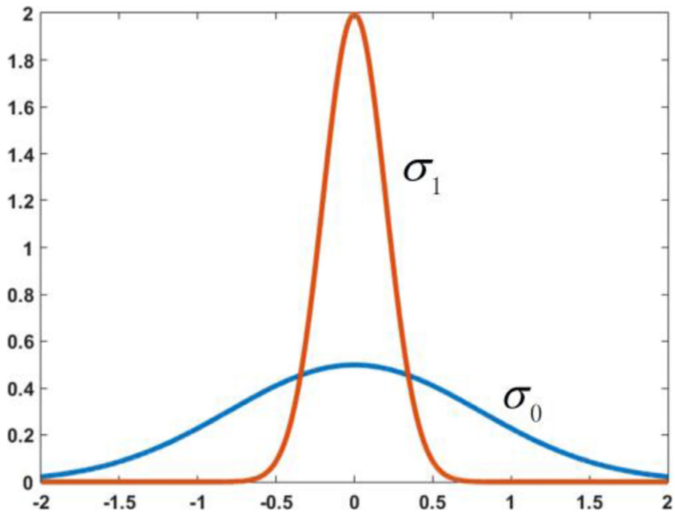


Fig. 4. The computed Gaussian distributions of dif_0 and dif_1.

Because Σ_I and Σ_E can be computed directly from sample mean and covariance of each class and all classes according to the computing method in [6], the total computations of Σ_I and Σ_E can be both reduced to $O(LNm^2)$. Adding the computation of feature transform, the computation of KRKISS requires $O(LNm^2 + dm^2)$ multiplication operations totally. Compared with the computation of KISS method ($O(Nd^2)$), the computation of KRKISS is even less when the dimension of original feature is much larger than the size of training samples. It is a common case in the applications of person re-identification and for example, in [6], the sizes of training samples in the expriments of VIPeR and QMUL dataset are 632 and 250, respectively, while the dimension of feature is 26,960 and much larger than the size of training samples. The efficientness of KRKISS is also validated through a related experiment in next section.

4. Experiments

In this section, we conduct experiments to evaluate the KRKISS method and compare it with the state of the art methods. All experiments were tested on publicly and standard datasets, namely VIPeR [25], PRID 450S [26] and CUHK01 [27]. For all the experiments, we used the cumulative matching characteristic (CMC) curve to evaluate the identification performance which is the most widely used evaluation methodology for person re-identification. It shows how performance improves as the number of requested images increases.

Fig. 6. Example images from VIPeR dataset [25].

The features used in our experiments includes LOMO [6], kCCA [5], FNN [17], SCNCD [4], ELF18 [19]. Moreover, we propose a novel feature which consists of different color, texture and shape feature. The extracted method of color feature is same as [5] and [29]. Then we extract the Scale Invariant Local Ternary Pattern (SILTP) [6] description as the texture feature. The shape feature is Histogram of Oriented Gradient (HOG) descriptor [23] and the extracting procedure is similar to [22]. We use the novel feature as original features of pedestrians in the following experiments unless noted otherwise.

4.1. Experiments on the VIPeR dataset

VIPeR [25] is a challenging person re-identification dataset which has been widely used for performance evaluation. It contains 632 pedestrian image pairs which are captured from two different camera views. The cameras have different viewpoints and illumination variations exist between them. All images have been scaled to 128 × 48 pixels. Fig 6 shows some example images from this dataset. In the experiments, the VIPeR dataset was randomly divided into half for training and the other half for testing. In

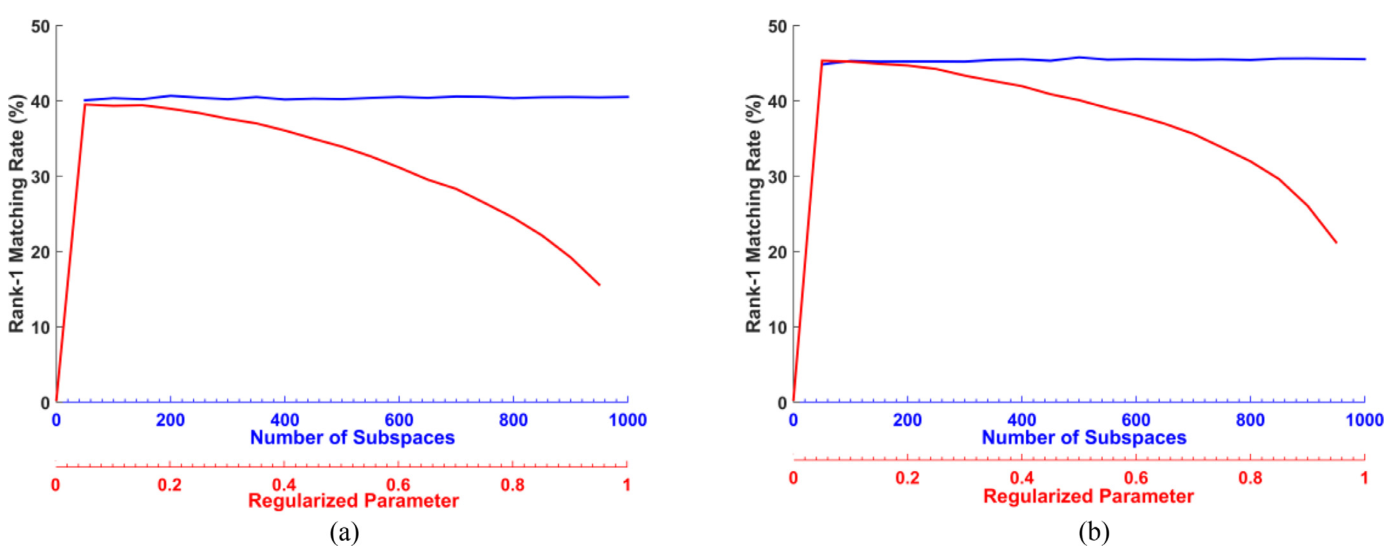


Fig. 5. CMC curves on the VIPeR dataset ($P = 316$). (a) Comparison of different parameters by using the LOMO feature. (b) Comparison of different parameters by using the kCCA feature.

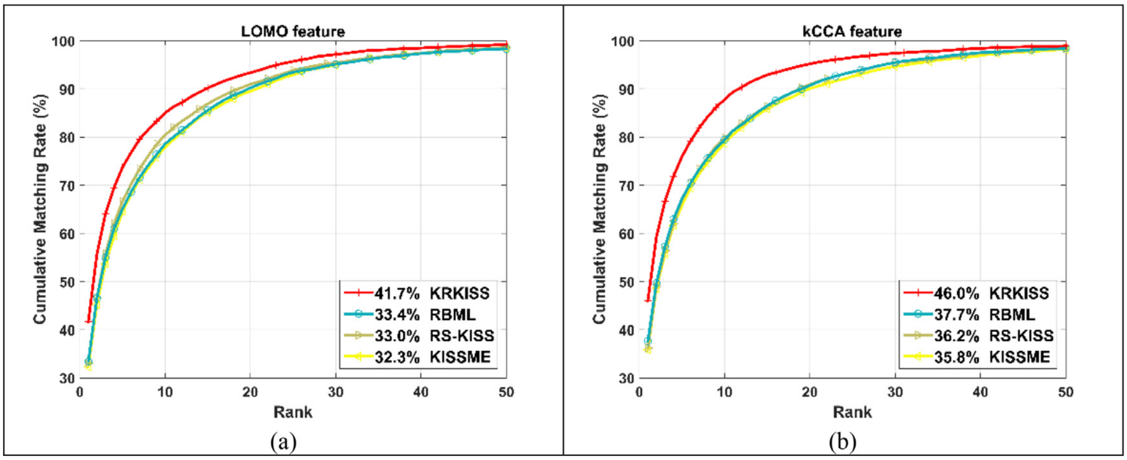


Fig. 7. CMC curves on the VIPeR dataset ($P = 316$). (a) Comparison of KISS based metric learning methods by using the LOMO feature. (b) Comparison of KISS based metric learning methods by using the kCCA feature.

Table 4
Comparison of different metric learning methods on the VIPeR database ($P = 316$). The cumulative matching scores (%) at rank 1, 5, and 10 are listed. Red and blue numbers are the best and second best results, respectively.

Feature/Method	LOMO($d = 26{,}960$)			kCCA($d = 5138$)			SCNCD($d = 70$)			ELF18($d = 8064$)			Reference
	$r = 1$	$r = 5$	$r = 10$	$r = 1$	$r = 5$	$r = 10$	$r = 1$	$r = 5$	$r = 10$	$r = 1$	$r = 5$	$r = 10$	
KRKISS	41.7	73.7	85.1	46.0	76.0	87.8	41.3	73.1	85.0	40.3	72.7	84.8	Ours
NFST [21]	40.1	70.3	82.3	44.9	74.5	86.0	36.7	67.6	80.0	37.8	70.3	84.0	2016 CVPR
MLAPG [16]	39.5	70.0	82.4	32.3	61.8	75.8	25.8	57.6	73.1	18.0	44.7	61.5	2015 ICCV
XQDA [6]	39.0	69.3	81.3	33.5	62.3	74.4	24.3	50.9	65.2	30.7	59.2	72.9	2015 CVPR
MFA [8]	35.4	67.9	81.0	40.4	70.1	83.1	37.9	68.9	81.4	36.8	70.3	83.0	2014 ECCV
kLFDA [8]	36.3	65.1	77.1	39.2	67.7	79.8	36.9	68.5	81.1	32.6	60.7	71.7	2014 ECCV
RBML [15]	33.4	65.1	78.6	37.7	67.2	79.5	37.7	65.2	77.6	33.9	63.8	77.1	2014 ECCV
LFDA [7]	37.6	68.7	81.3	34.8	64.6	77.5	16.3	29.8	36.6	33.1	63.3	76.8	2013 CVPR
RS-KISS [13]	33.0	66.7	80.4	36.2	67.0	79.8	37.1	67.2	79.7	34.2	63.6	76.6	2013 IEEE
KISSME [12]	32.3	64.6	77.9	35.8	65.8	78.8	38.7	69.8	82.8	32.7	63.5	78.5	2012 CVPR

order to obtain a precise experimental result, this procedure was repeated 10 times and an average value was taken.

4.1.1. Comparison of metric learning algorithms

Firstly, in order to demonstrate the improvement to KISS method, we compare our KRKISS method with three KISS based methods (KISSME [12], RS-KISS [13] and RBML [15]) by using two features (LOMO [6] and the kCCA [5]) respectively. Both features have greatly identification ability for person re-identification. The experimental results are shown in Fig 7. In the experiment of the LOMO feature, it can be seen that our proposed method (KRKISS) is obviously better than other KISS based methods, achieving a rate of 41.7% at rank = 1. In the experiments of the kCCA feature, our proposed KRKISS method improves the rank-1 accuracy from 37.7% to 46.0%. Both experimental results show the greatly improvement of KRKISS method to KISS based methods and demonstrate that the kernelized feature transform and the random subspace ensemble method are necessity for KISS based methods.

Next we compare the KRKISS method with some state-of-the-art metric learning methods which were reported having good performance on the VIPeR dataset. The compared metric learning methods includes NFST [21], MLAPG [16], XQDA [6], MFA [8], kLFDA [8], RBML [15], LFDA [7], RS-KISS [13] and KISSME [12]. For testing the robustness of metric learning methods to different features, we apply four kinds of features (LOMO [6], kCCA [5], SCNCD [4] and ELF18 [19]) for all metric learning methods in the experiments. The experimental results are shown in Table 4 and Fig 8. It can be seen that our proposed KRKISS method achieves the best matching rates in all experiments with four features, which demonstrates that the KRKISS method not noly has better performance than all the compared methods, but also has good stability and strong robustness to different kinds of features. Meanwhile, the average training time comparison of metric learning methods is shown in Table 5 (including subspace learning time). The training was performed on a notebook PC with an Intel i5-3210 @2.50 GHz CPU. Table 5 shows that the KRKISS is efficient and the time consumption is even less than KISSME [12] method. The theoretical analysis and the experimental results both show the efficientness of our proposed KRKISS method. Although it needs to compute several Mahalanobis matrices in different subspaces, the KRKISS method is still efficient because the dimension of subspace is greatly less than original features.

Table 5
Training time (seconds) of metric learning algorithms.

	KRKISS	NFST	MLAPG	XQDA	MFA	kLFDA	KISSME	LFDA
Time	5.04	2.48	40.9	3.86	2.58	2.74	7.41	229.3

4.1.2. Comparison to the state of the art

In order to demonstrate that our proposed method (KRKISS) can achieve state-of-the-art performance for person re-identification, we compare it with other state-of-the-art methods like GOG [29], FNN [17], NFST [21], etc. The compared methods include two feature represent based methods [6,29] and five metric learning based methods [6,16,21]. In the experiment, all the parameter settings, the used features and the metric learning method were refer to their own paper. The experimental result is shown in Table 6 and Fig 9. It can be seen that our proposed method (KRKISS) achieves the best performance while it improves the most

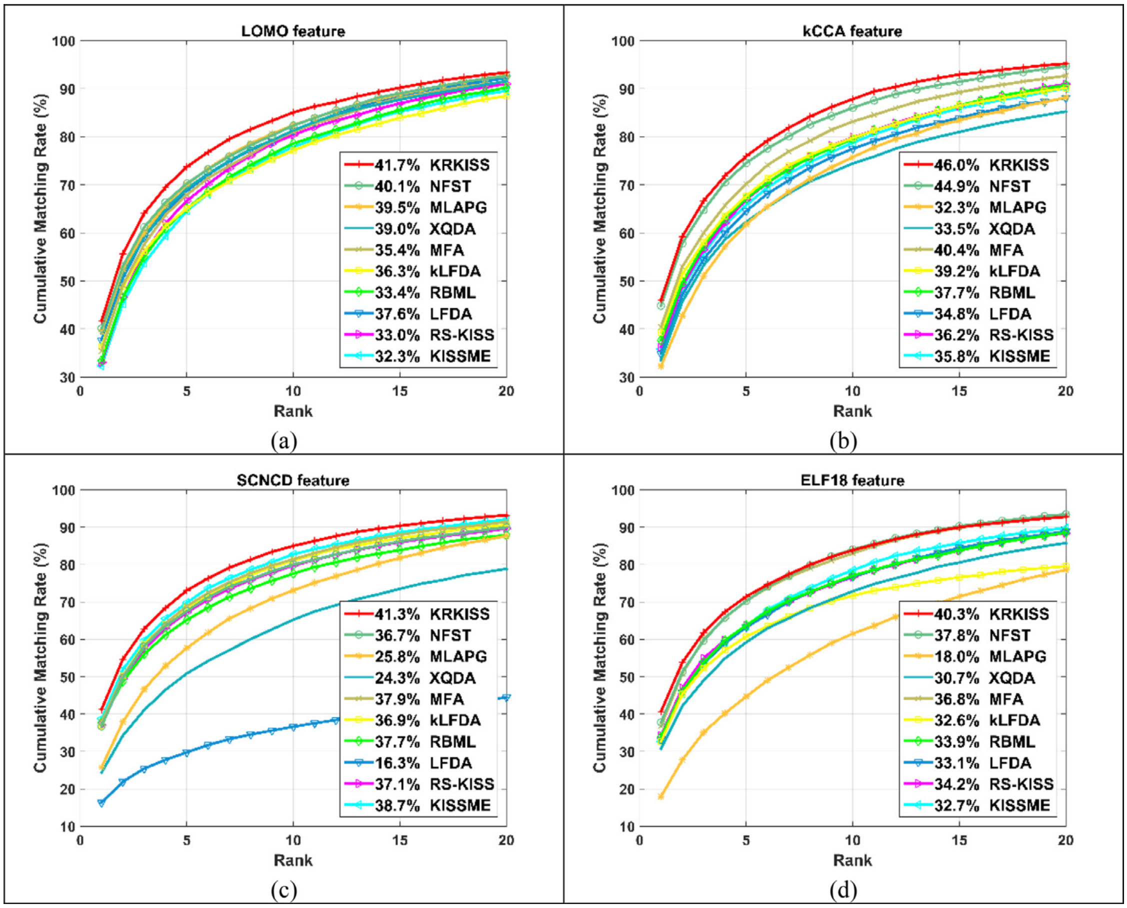


Fig. 8. CMC curves on the VIPeR dataset (P = 316). (a) Comparison of different metric learning methods by using the LOMO feature. (b) Comparison of different metric learning methods by using the kCCA feature. (c) Comparison of different metric learning methods by using the SCNCD feature. (b) Comparison of different metric learning methods by using the ELF18 feature.

Table 6
Comparison of state-of-the-art results reported with the VIPeR database ($P = 316$). The cumulative matching scores (%) at rank 1, 5, 10, and 20 are listed. The best and the second best scores are respectively shown in red and blue.

Method	$r = 1$	$r = 5$	$r = 10$	$r = 20$	Reference
KRKISS	54.2	83.5	91.3	96.4	Ours
GOG [29]	47.0	75.1	85.4	93.9	2016 CVPR
NFST [21]	42.3	71.5	82.9	92.1	2016 CVPR
MLAPG [16]	39.5	70.0	82.4	92.8	2015 ICCV
LOMO + XQDA [6]	39.0	69.3	81.3	91.6	2015 CVPR
RBML [15]	33.4	65.1	78.6	90.2	2014 ECCV
RS-KISS [13]	33.0	66.7	80.4	91.0	2013 IEEE
KISSME [12]	32.3	64.6	77.9	89.5	2012 CVPR

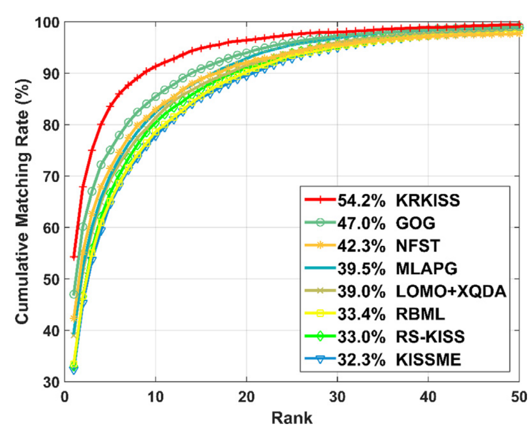


Fig. 9. The CMC curves and rank-1 matching rates on the VIPeR dataset.

important index for person re-identification (rank-1 accuracy) from 47.0% to 54.2%. It proves our model can effectively improve the performance of person *Re*-ID.

4.2. Experiments on the PRID 450S dataset

In this section, we conduct some experiments to evaluate the KRKISS method on the PRID 450S dataset [26]. The PRID 450S dataset consists of 450 images pairs of pedestrians with significant differences in background, viewpoint and illumination. It has been recently released and is regarded as more realistic. Some example images from the PRID 450S dataset are shown in Fig 10. In the experiments, the PRID 450S dataset was randomly divided into half for training and half for testing and the procedure was repeated 10 times to record an average performance.

4.2.1. Comparison of metric learning algorithms

Firstly, we compare our proposed KRKISS method with other three KISS based methods (KISSME [12], RS-KISS [13] and RBML [15]). The used features include two efficient features, LOMO [6] and FFN [17]. Among them, the LOMO feature is artificially designed and the FNN feature is learned by a deep learning method. The experimental results are shown in Fig 11. As we can see in Fig 11, whether using the LOMO feature or the FNN feature, the red line which shows the cumulative matching rate curve of KRKISS method is higher than other lines. It demonstrates

Table 7
Comparison of different metric learning methods on the PRID 450S database ($P = 225$). The cumulative matching scores (%) at rank 1, 5 and 10 are listed. The best and the second best scores are respectively shown in red and blue.

Method/Feature	LOMO($d = 26{,}960$)			FFN($d = 4096$)			SCNCD($d = 70$)			ELF18($d = 8064$)			
	$r = 1$	$r = 5$	$r = 10$	$r = 1$	$r = 5$	$r = 10$	$r = 1$	$r = 5$	$r = 10$	$r = 1$	$r = 5$	$r = 10$	Reference
KRKISS	58.0	83.8	90.9	53.0	77.5	85.8	43.7	71.0	81.6	31.9	59.5	72.3	Ours
NFST [21]	57.1	82.8	89.8	54.4	79.0	86.9	32.0	57.7	68.6	29.5	54.1	67.4	2016 CVPR
MLAPG [16]	55.5	80.1	88.3	22.7	46.5	57.9	30.4	60.3	73.6	8.8	23.8	35.2	2015 ICCV
XQDA [6]	55.9	80.6	88.0	48.4	71.6	80.9	16.1	40.7	53.6	27.8	51.5	63.0	2015 CVPR
MFA [8]	50.8	78.7	87.7	47.9	75.1	84.6	41.6	70.5	80.6	30.0	57.2	69.7	2014 ECCV
kLFDA [8]	51.3	78.8	88.2	48.6	75.7	84.9	20.9	39.3	49.4	9.1	23.9	32.5	2014 ECCV
RBML [15]	54.0	80.7	88.6	51.3	75.9	84.1	35.6	63.0	74.3	29.8	55.3	67.0	2014 ECCV
LFDA [7]	54.1	79.4	87.8	49.9	75.0	83.8	28.9	50.9	62.1	30.5	55.1	66.9	2013 CVPR
RS-KISS [13]	54.6	79.5	88.1	49.5	74.4	83.6	33.4	61.1	72.4	29.1	52.9	65.4	2013 IEEE
KISSME [12]	47.1	75.5	85.6	41.8	69.4	79.7	33.9	58.9	69.3	25.9	53.6	66.7	2012 CVPR

Fig. 10. Example images from PRID 450S dataset [25].

that our proposed KRKISS method has the highest performance in all compared KISS-based methods. Meanwhile, it further proves the necessity of kernelized feature transform and the random subspace ensemble for KISS based methods.

Next, we compare our proposed KRKISS method with some state-of-the-art metric learning methods [6,7,8,12,13, 15,16,21]. The features we used in this experiment include LOMO [6], FFN [17], SCNCD [4], and ELF18 [19] which were reported with good performance on the PRID 450S dataset. The experimental results are shown in Table 7 and Fig 12. It can be seen that the KRKISS method achieves the best performance with three features (LOMO, SCNCD, ELF18) and the second best with FNN feature. The results of comparing with other state-of-the-art metric learning methods demonstrate the effectiveness and the robustness of our proposed KRKISS algorithm.

Table 8
Comparison of state-of-the-art results reported with the PRID 450S database ($P = 225$). The cumulative matching scores (%) at rank 1, 5, 10, and 20 are listed. The best and the second best scores are respectively shown in red and blue.

Method	$r = 1$	$r = 5$	$r = 10$	$r = 20$	Reference
KRKISS	70.4	89.0	94.1	97.8	Ours
GOG [29]	66.4	87.6	93.8	97.5	2016 CVPR
NFST [21]	68.0	87.0	92.6	96.4	2016 CVPR
MLAPG [16]	55.5	80.1	88.3	94.7	2015 ICCV
LOMO + XQDA [6]	55.9	80.6	88.0	93.4	2015 CVPR
RBML [15]	54.0	80.7	88.6	93.9	2014 ECCV
RS-KISS [13]	54.6	79.5	88.1	94.0	2013 IEEE
KISSME [12]	46.4	74.9	84.9	92.3	2012 CVPR

4.2.2. Comparison to the state of the art

Next we compare our proposed method with the state-of-the-art methods on the PRID 450S dataset. The compared methods and the experimental setups are same as to the experiments on the VIPeR dataset. The cumulative matching rates at rank 1, 5, 10, and 20 are listed in Table 8 and the CMC curves are shown in Fig 13. As we can see in Table 8, the rank-1, rank-5, rank-10 and rank-20 matching rates of our proposed method achieve 70.4%, 89.0%, 94.1% and 97.7%, respectively, which are the highest in all compared methods. The PRID 450S dataset is regarded as more realistic for person re-identification, and thus the experimental results on the PRID 450S dataset demonstrate that our proposed method is robust to different environments of person *Re*-ID.

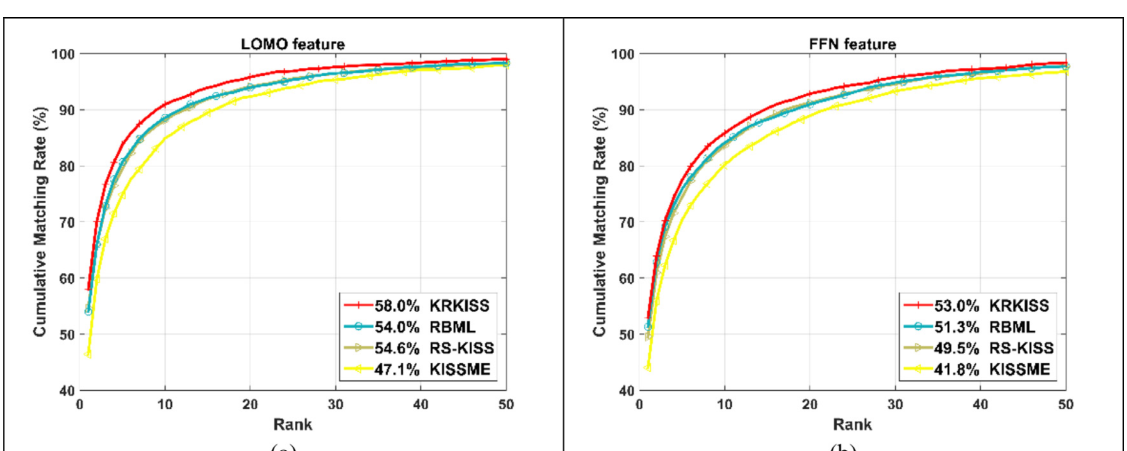


Fig. 11. CMC curves on the PRID 450S dataset ($P = 225$). (a) Comparison of KISS based metric learning methods by using the LOMO feature. (b) Comparison of KISS based metric learning methods by using the FFN feature.

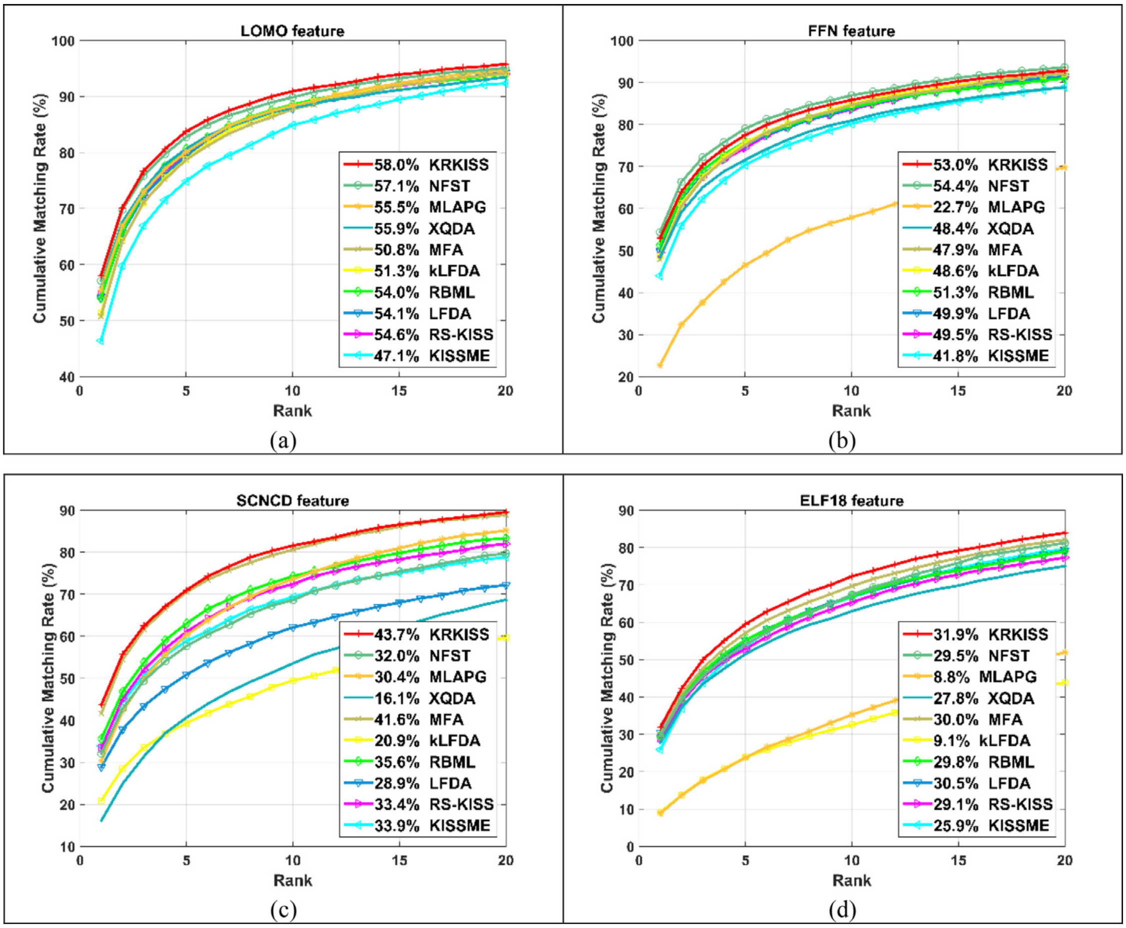


Fig. 12. CMC curves on the PRID 450S dataset ($P = 225$). (a) Comparison of different metric learning methods by using the LOMO feature. (b) Comparison of different metric learning methods by using the FFN feature. (c) Comparison of different metric learning methods by using the SCNCD feature. (b) Comparison of different metric learning methods by using the ELF18 feature.

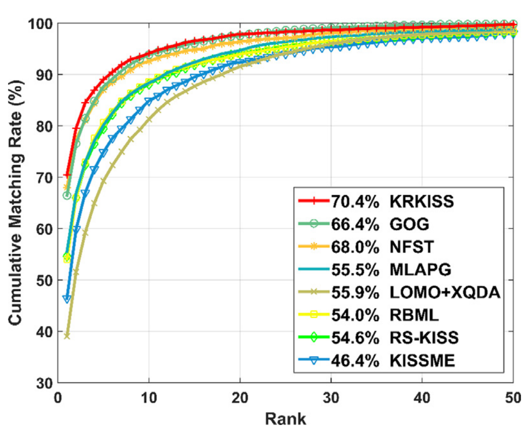


Fig. 13. The CMC curves and rank-1 matching rates on the PRID 450S dataset.

4.3. Experiments on the CUHK01 dataset

CUHK01 [27] is another challenging person re-identification dataset which contains 971 pedestrians. Each pedestrian has four images captured from two different camera views in a campus environment. Camera A captures the frontal view or back view of people while camera B captures the side view. Large inter-camera variations are observed in this dataset which makes person re-identification challenging. All images have been normalized to 160×60 pixels. Fig 14 shows some example images from this dataset. Different with the experimental settings of VIPeR dataset and PRID 450S dataset, the expriments of CUHK01 dataset can be divided into CUHK01(M = 1) and CUHK01(M = 2) because each pedestrian has more than two images. The experimental setting of

Fig. 14. Example images from CUHK01 dataset [27].

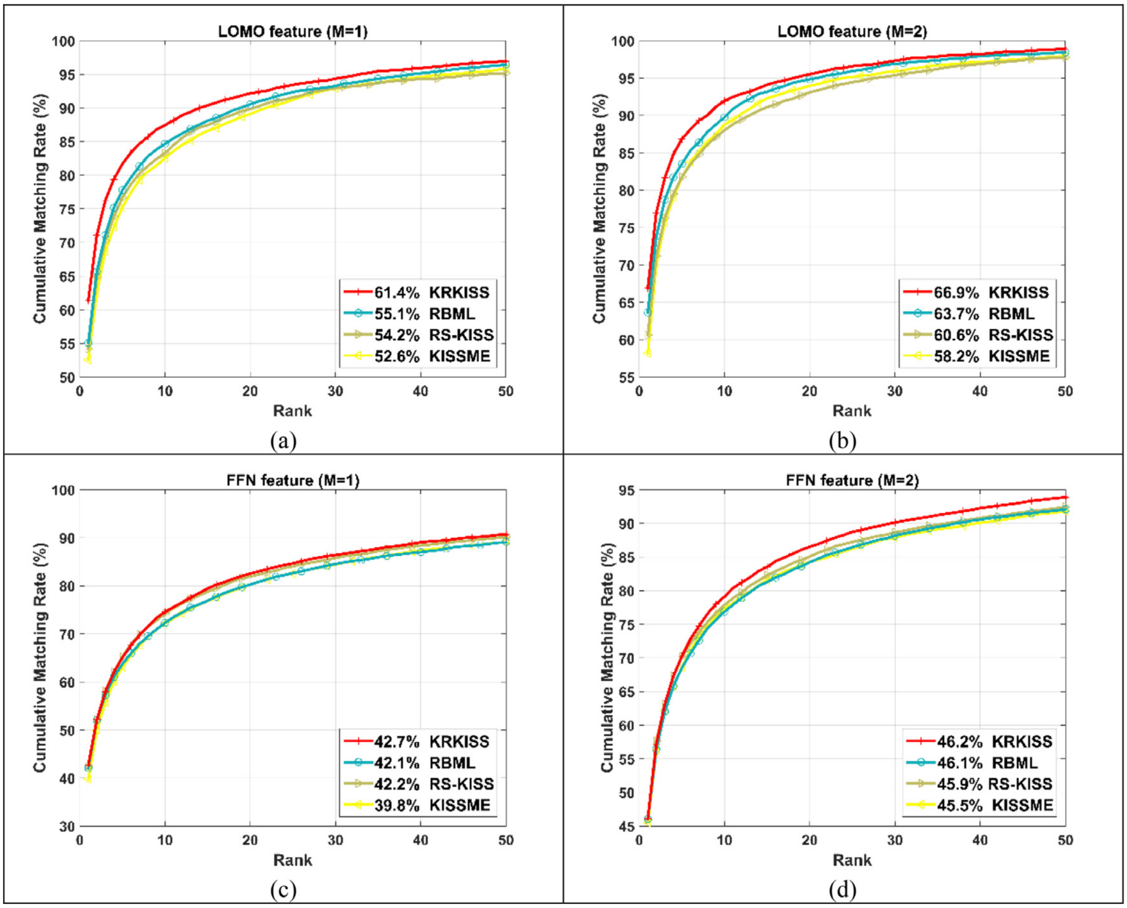


Fig. 15. CMC curves on the CUHK01 dataset ($P = 486$). (a) Comparison of KISS based metric learning methods with the single shot setting by using the LOMO feature. (b) Comparison of KISS based metric learning methods with the multi shot setting by using the LOMO feature. (c) Comparison of KISS based metric learning methods with the single shot setting by using the FFN feature. (d) Comparison of KISS based metric learning methods with the multi shot setting by using the FFN feature.

Table 9
Comparison of different metric learning methods on the CUHK01 database ($P = 486$). The cumulative matching scores (%) at rank 1, 5 and 10 are listed. The best and the second best scores are respectively shown in red and blue.

Feature/Method	$M = 1$						$M = 2$						Reference
	LOMO($d = 26{,}960$)			FFN($d = 4096$)			LOMO($d = 26{,}960$)			FFN($d = 4096$)			
	$r = 1$	$r = 5$	$r = 10$	$r = 1$	$r = 5$	$r = 10$	$r = 1$	$r = 5$	$r = 10$	$r = 1$	$r = 5$	$r = 10$	
KRKISS	61.4	81.7	87.4	42.7	65.2	74.6	66.9	86.9	91.9	46.2	70.3	79.1	Ours
NFST [21]	60.0	80.8	86.7	39.4	60.6	69.0	65.0	85.0	89.9	43.4	66.5	74.4	2016 CVPR
MLAPG [16]	58.4	79.0	85.5	24.8	46.5	56.9	64.7	86.6	91.6	26.8	50.3	61.3	2015 ICCV
XQDA [6]	55.8	78.6	85.7	34.5	55.6	63.9	62.8	83.9	90.5	39.7	60.1	68.4	2015 CVPR
MFA [8]	58.7	81.2	88.1	39.6	64.2	73.9	63.8	85.7	91.9	44.7	70.6	79.9	2014 ECCV
kLFDA [8]	58.9	80.8	86.9	32.0	53.8	64.2	64.7	84.0	90.4	34.1	58.5	69.0	2014 ECCV
RBML [15]	55.1	77.8	84.7	42.1	63.7	72.3	63.7	83.5	89.8	46.1	68.5	76.9	2014 ECCV
LFDA [7]	55.6	77.9	85.1	39.3	63.6	73.1	62.9	83.5	89.7	44.8	68.7	77.7	2013 CVPR
RS-KISS [13]	54.2	76.7	83.3	42.2	65.2	74.3	60.6	81.8	88.0	45.9	70.2	77.9	2013 IEEE
KISSME [12]	52.6	75.2	82.5	39.8	63.1	72.1	58.2	81.7	88.8	45.5	68.6	77.3	2012 CVPR

CUHK01(M = 1) is the single shot setting and CUHK01(M = 2) is the multi shot setting, which are common to [46] and [6] respectively.

4.3.1. Comparison of metric learning algorithms

Same as to the experiments in PRID 450S dataset, we compare our proposed KRKISS method with other three KISS based methods (KISSME [12], RS-KISS [13] and RBML [15]) by using two features (LOMO [6] and the FFN [17]) respectively. The experimental matching rate curves are shown in Fig 15. Meanwhile, we have performed some experiments to compare the KRKISS method with some state-of-the-art metric learning methods. The experimental results are shown in Table 9 and Fig 16.

As we can see in Fig 15, our proposed KRKISS method achieves better matching rates than other three KISS based methods no matter the experimental setting is the single shot setting or the multi shot setting. In Table 9 and Fig 16, compared with other state-of-the-art metric learning methods, the KRKISS method achieves the best rank-1 rate when the experimental setting is the single shot setting, and it is among the top two when the experimental setting is the multi shot setting. All the experimental results indicates that the KRKISS metric learning method is not only suit for the single-shot

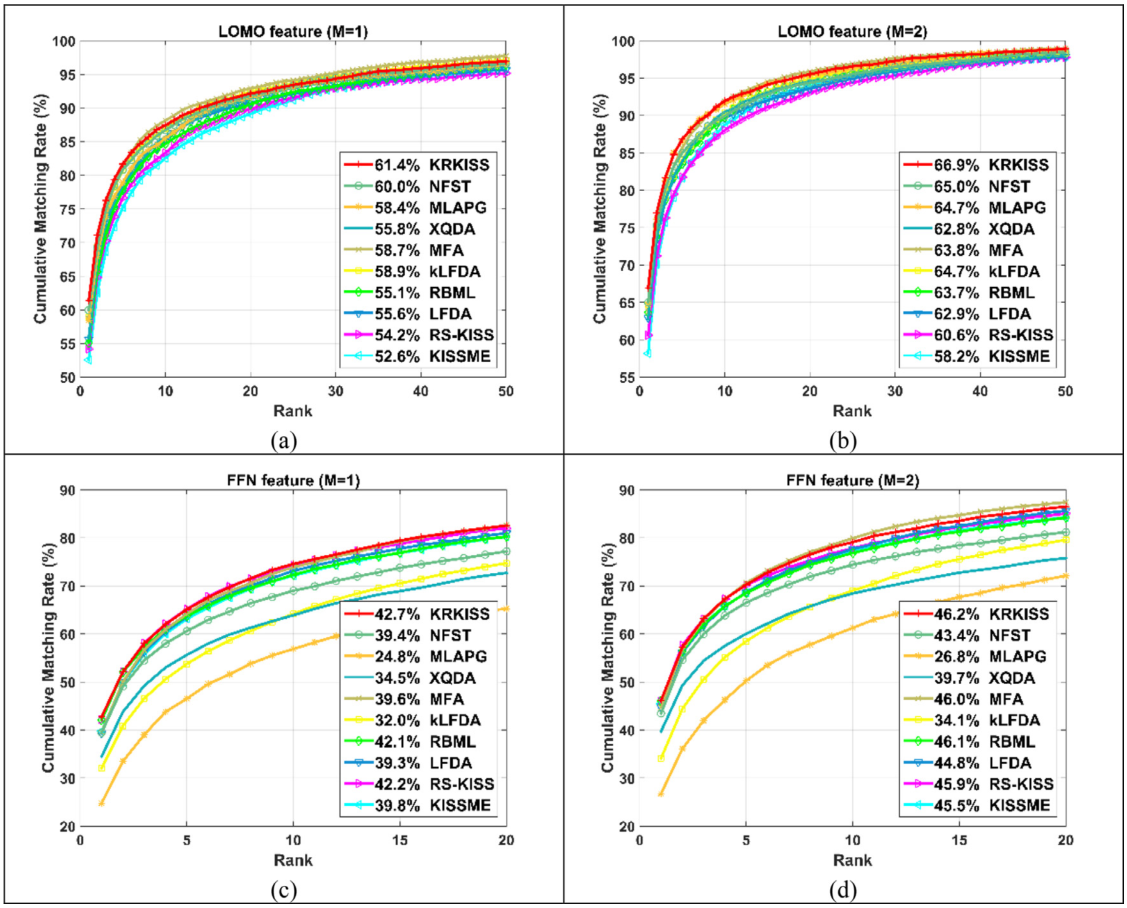


Fig. 16. CMC curves on the CUHK01 dataset ($P = 486$). (a) Comparison of different metric learning methods with the single shot setting by using the LOMO feature. (b) Comparison of different metric learning methods with the multi shot setting by using the LOMO feature. (c) Comparison of different metric learning methods with the single shot setting by using the FFN feature. (b) Comparison of different metric learning methods with the multi shot setting by using the FFN feature.

Table 10
Comparison of state-of-the-art results reported with the CUHK01 database ($P = 486$). The cumulative matching scores (%) at rank 1, 5, 10, and 20 are listed. The best and the second best scores are respectively shown in red and blue.

Method	CUHK01(M = 1)				CUHK01(M = 2)				Reference
	r = 1	r = 5	r = 10	r = 20	r = 1	r = 5	r = 10	r = 20	
KRKISS	66.9	83.0	88.7	92.7	72.9	89.8	94.0	96.9	Ours
NFST [21]	60.0	80.8	86.7	91.9	65.0	85.0	89.9	94.4	2016 CVPR
GOG [29]	59.8	77.6	84.3	89.9	66.4	84.9	90.4	94.5	2016 CVPR
MLAPG [16]	58.4	79.0	85.5	91.6	64.7	86.6	91.6	95.4	2015 ICCV
LOMO + XQDA [6]	55.8	78.6	85.7	91.3	62.8	83.9	90.5	94.6	2015 CVPR
RBML [15]	55.1	77.8	84.7	90.5	63.6	83.5	89.8	94.8	2014 ECCV
RS-KISS [13]	54.2	76.7	83.3	89.8	60.6	81.8	88.0	93.1	2013 IEEE
KISSME [12]	52.6	75.2	82.5	89.4	58.2	81.7	88.8	94.0	2012 CVPR

cases in person re-identification, but also suit for the multi-shot cases.

4.3.2. Comparison to the state of the art

Next we compare our proposed methods (KRKISS) with some state-of-the-art methods which were reported with good performance on the CUHK01 dataset. The compared methods and the experimental setups are same as to the experiments on the VIPeR dataset. The experimental results are shown in Table 10 and Fig 17. It can be observed that our proposed method achieves the best matching rates on CUHK01 (M = 1) and CUHK01(M = 2), which are 6.9% and 7.9% better than the second at rank-1 respectively. The significantly improvement of performance demonstrates that our proposed KRKISS method is also suitable to multi shot scenes in the applications of person *Re*-ID.

5. Conclusions and future works

In this paper, we have presented a novel metric learning method for person re-identification called Kernelized Random KISS (KRKISS). The KRKISS method can augment Guassian distribution of data by transforming the original features into kernelized feature. By doing so, the differences between kernelized features pairs can better fit to Gaussian distribution, and thus they can be more suitable for the Guassian assumption based models. Meanwhile, the SSS problem in covariance matrix estimation based methods can now be solved by a random subspace ensemble method because the influence of SSS problem would decreases or disappears when the dimension of subspace is greatly less than number of samples. The experiments on three publicly datasets, VIPeR, PRID 450S, CUHK01, show the effectiveness and the robust-

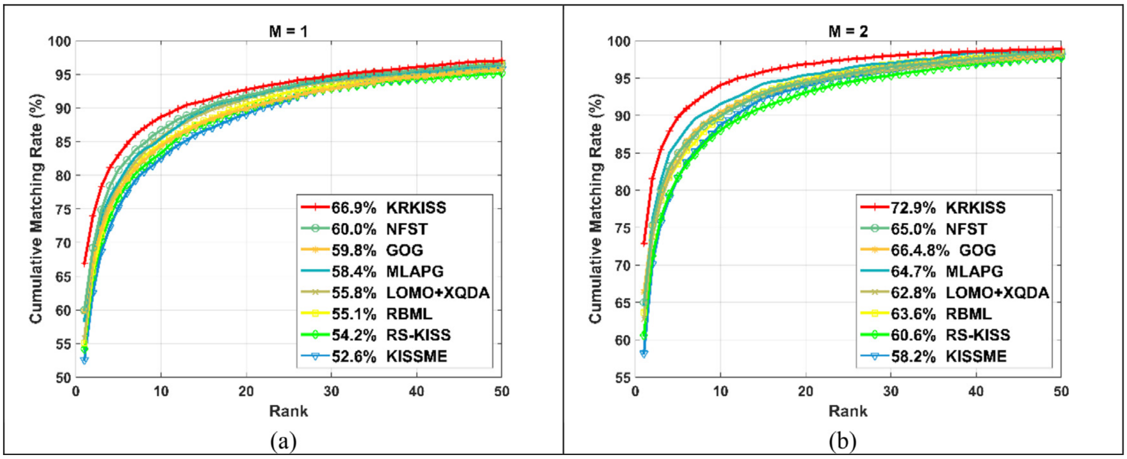


Fig. 17. The CMC curves and rank-1 matching rates on the CUHK01 dataset. (a) Comparison of some state-of-the-art methods with the single shot setting. (b) Comparison of some state-of-the-art methods with the multi shot setting.

ness of KRKISS method, and demonstrate that the KRKISS method is better than the state-of-the-art metric learning methods in most cases. It would be interesting to see that the KRKISS method can apply to other cross-view matching problem, such as face recognition.

Acknowledgements

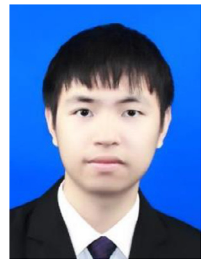
The authors would like to thank the anonymous reviewers for their critical and constructive comments and suggestions. This work is supported by China National Natural Science Foundation under Grant nos. 61673299, 61203247 and The Hong Kong Polytechnic University (Project account: G-UC42). This work is also partially supported by China National Natural Science Foundation under Grant nos. 61573259, 61573255, 61375012. It is also supported by the Fundamental Research Funds for the Central Universities (Grant No. 0800219327). It is also partially supported by Fujian Provincial Key Laboratory of Information Processing and Intelligent Control (Minjiang University) and by the Open Project Program of Key Laboratory of Intelligent Perception and Systems for High-Dimensional Information of Ministry of Education under Grant no. 30920130122005.

Reference

[1] R Vezzani, D Baltieri, R Cucchiara, People reidentification in surveillance and forensics: a survey, ACM Comput. Surv. (CSUR) 46 (2) (2013) 29.
[2] A Bedagkar-Gala, K Shah S, A survey of approaches and trends in person re-identification, Image Vis. Comput. 32 (4) (2014) 270–286.
[3] D Gray, H Tao, Viewpoint invariant pedestrian recognition with an ensemble of localized features, in: Proceedings of European conference on computer vision, Springer Berlin Heidelberg, 2008, pp. 262–275.
[4] Y Yang, J Yang, J Yan, Salient color names for person re-identification, in: Proceedings of European Conference on Computer Vision, Springer International Publishing, 2014, pp. 536–551.
[5] G Lisanti, I Masi, A Del Bimbo, Matching people across camera views using kernel canonical correlation analysis, in: Proceedings of the International Conference on Distributed Smart Cameras, ACM, 2014, p. 10.
[6] S Liao, Y Hu, X Zhu, Person re-identification by local maximal occurrence representation and metric learning, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 2197–2206.
[7] S Pedagadi, J Orwell, S Velastin, Local fisher discriminant analysis for pedestrian re-identification, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2013, pp. 3318–3325.
[8] F Xiong, M Gou, O Camps, Person re-identification using kernel-based metric learning methods, in: Proceedings of European Conference on Computer Vision, Springer International Publishing, 2014, pp. 1–16.
[9] A Mignon, F Jurie, Pcca: A new approach for distance learning from sparse pairwise constraints, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2012, IEEE, 2012, pp. 2666–2672.
[10] Q Weinberger K, K Saul L, Distance metric learning for large margin nearest neighbor classification, J. Mach. Learn. Res. 10 (Feb) (2009) 207–244.
[11] Z Li, S Chang, F Liang, Learning locally-adaptive decision functions for person verification, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2013, pp. 3610–3617.
[12] M Köstinger, M Hirzer, P Wohlhart, Large scale metric learning from equivalence constraints, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2012, pp. 2288–2295.
[13] D Tao, L Jin, Y Wang, Person re-identification by regularized smoothing kiss metric learning, in: Proceedings of IEEE Transactions on Circuits and Systems for Video Technology, 23, 2013, pp. 1675–1685.
[14] D Tao, L Jin, Y Wang, Person reidentification by minimum classification error-based KISS metric learning, IEEE Trans. Cybern. 45 (2) (2015) 242–252.
[15] E Liong V, J Lu, Y Ge, Regularized Bayesian metric learning for person re-identification, in: Proceedings of European Conference on Computer Vision, Springer International Publishing, 2014, pp. 209–224.
[16] S Liao, Z Li S, Efficient psd constrained asymmetric metric learning for person re-identification, in: Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 3685–3693.
[17] S Wu, C Chen Y, X Li, An enhanced deep feature representation for person re-identification, in: Proceedings of 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), IEEE, 2016, pp. 1–8.
[18] M Farenzena, L Bazzani, A Perina, Person re-identification by symmetry-driven accumulation of local features, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2010, IEEE, 2010, pp. 2360–2367.
[19] C Chen Y, S Zheng W, J Lai, Mirror representation for modeling view-specific transform in person re-identification, Proc. IJCAI (2015) 3402–3408.
[20] B Ma, Y Su, F Jurie, Local descriptors encoded by fisher vectors for person re-identification, in: Proceedings of European Conference on Computer Vision, Springer Berlin Heidelberg, 2012, pp. 413–422.
[21] Zhang L, Xiang T, Gong S. Learning a discriminative null space for person re-identification. Proceedings of the IEEE Conference on Computer Vision and Pattern Recogni-tion. 2016: 12391248.
[22] G Lisanti, I. Masi, AD. Bagdanov, Person re-identification by iterative re-weighted sparse ranking, IEEE Trans. Pattern Anal. Mach. Intell. 37 (8) (2015) 1629–1642.
[23] N Dalal, B Triggs, Histograms of oriented gradients for human detection, in: Proceedings of IEEE IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 1, 2005, pp. 886–893.
[24] T Ahonen, A Hadid, M Pietikainen, Face description with local binary patterns: Application to face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 28 (12) (2006) 2037–2041.
[25] Gray D, Brennan S, Tao H. Evaluating appearance models for recognition, reacquisition, and tracking. Proceedings of IEEE International Workshop on Performance Evaluation for Tracking and Surveillance (PETS). 2007, 3(5).
[26] M Roth P, M Hirzer, M Koestinger, Mahalanobis distance learning for person re-identification, Person Re-Identification, Springer, London, 2014, pp. 247–267.
[27] W Li, R Zhao, X Wang, Human reidentification with transferred metric learning, in: Proceedings of Asian Conference on Computer Vision, Springer, Berlin Heidelberg, 2012, pp. 31–44.
[28] W Li, R Zhao, T Xiao, Deepreid: deep filter pairing neural network for person re-identification, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 152–159.
[29] T Matsukawa, T Okabe, E Suzuki, Hierarchical gaussian descriptor for person re-identification, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 1363–1372.
[30] C Jose, F Fleuret, Scalable metric learning via weighted approximate rank component analysis, European conference on computer vision (2016) 875–890.

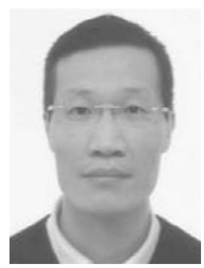
[31] R Zhao, W Ouyang, X Wang, Person re-identification by salience matching, in: Proceedings of the IEEE International Conference on Computer Vision, 2013, pp. 2528–2535.
[32] Z Shi, M Hospedales T, T Xiang, Transferring a semantic representation for person re-identification and search, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 4184–4193.
[33] Y Shen, W Lin, J Yan, Person re-identification with correspondence structure learning, in: Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 3200–3208.
[34] R Zhao, W Ouyang, X Wang, Learning mid-level filters for person re-identification, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 144–151.
[35] B Schölkopf, J Smola A, Learning With Kernels: Support Vector Machines, Regularization, Optimization, and Beyond, MIT press, 2002.
[36] D Tao, Y Guo, M Song, Person re-identification by dual-regularized kiss metric learning, IEEE Trans. Image Process. 25 (6) (2016) 2726–2738.
[37] JV. Davis, B. Kulis, P. Jain, Information-theoretic metric learning, in: Proceedings of the 24th international conference on Machine learning, ACM, 2007, pp. 209–216.
[38] W Zhao, R Chellappa, PJ. Phillips, Face recognition: A literature survey, ACM Comput. Surv. (CSUR) 35 (4) (2003) 399–458.
[39] L Lee, WEL. Grimson, Gait analysis for recognition and classification, in: Proceedings.of the Fifth IEEE International Conference on. IEEE, 2002, 2002, pp. 148–155.
[40] J. Daugman, How iris recognition works, IEEE Trans. Circuits Syst. Video Technol. 14 (1) (2004) 21–30.
[41] W Wang, A Taalimi, K Duan, Learning patch-dependent kernel forest for person re-identification, in: Proceedings of IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, 2016, pp. 1–9.
[42] G Lisanti, S Karaman, I Masi, Multi channel-kernel ca-nonical correlation analysis for cross-view person re-identification, ACM Transactions on Multimedia Computing, Communications, and Applications 13 (2) (2017) 2651–2664.
[43] H Liu, M Qi, J Jiang, Kernelized relaxed margin components analysis for person re-identification, IEEE Signal Process. Lett. 22 (7) (2015) 910–914.
[44] T Xiao, H Li, Ouyang W, Learning deep feature representations with domain guided dropout for person re-identification, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 1249–1258.
[45] HW. Lilliefors, On the Kolmogorov-Smirnov test for normality with mean and variance unknown, J. Am. Stat. Assoc. 62 (318) (1967) 399–402.
[46] S Paisitkriangkrai, C Shen, A van den Hengel, Learning to rank in person re-identification with metric ensembles, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1846–1855.
[47] Y Duan, J Lu, J Feng, Deep localized metric learning, IEEE Trans. Circuits Syst. Video Technol. 99 (2017) In press, doi:10.1109/TCSVT.2017.2711015.
[48] J Hu, J Lu, Y.P. Tan, in: Deep transfer metric learning, 2015, pp. 325–333.
[49] C Stanfill, D Waltz, Toward memory-based reasoning, Commun. ACM 29 (12) (1986) 1213–1228.
[50] C Li, H Li, One dependence value difference metric, Knowl.-Based Syst. 24 (5) (2011) 589–594.
[51] L Jiang, C Li, An augmented value difference measure, Pattern Recogn. Lett. 34 (10) (2013) 1169–1174.
[52] L Jiang, C Li, H Zhang, A novel distance function: Frequency difference metric, Int. J. Pattern Recogn. Artif. Intell. 28 (02) (2014) 1451002.
[53] C Li, L Jiang, H Li, Toward value difference metric with attribute weighting, Knowl. Inf. Syst. 50 (3) (2017) 795–825.
[54] DR Wilson, TR. Martinez, Improved heterogeneous distance functions, J. Artif. Intell. Res. 6 (1) (1997) 1–34.
[55] C Li, L Jiang, H Li, Naive Bayes for value difference metric, Front. Comput. Sci. 8 (2) (2014) 255–264.
[56] C Li, L Jiang, H Li, Naive Bayes for value difference metric, Front. Comput. Sci. 8 (2) (2014) 255–264.
[57] R Short, K Fukunaga, The optimal distance measure for nearest neighbor classification, IEEE Trans. Inf. Theory 27 (5) (1981) 622–627.
[58] C Li, L Jiang, H Li, Local value difference metric, Pattern Recogn. Lett. 49 (2014) 62–68.
[59] F Shen, C Shen, Q Shi, Hashing on nonlinear manifolds, IEEE Trans. Image Process. 24 (6) (2015) 1839–1851.
[60] F Shen, X Zhou, Y Yang, A fast optimization method for general binary code learning, IEEE Trans. Image Process. 25 (12) (2016) 5610–5621.

Cairong Zhao is currently an associate professor at Tongji University. He received the Ph.D. degree from Nanjing University of Science and Technology, M.S. degree from Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, and B.S. degree from Jilin University, in 2011, 2006 and 2003, respectively. He is the author of more than 30 scientific papers in pattern recognition, computer vision and related areas. His research interests include computer vision, pattern recognition and visual surveillance.



Yipeng Chen is currently a master candidate in College of Electronics and Information Engineering, Tongji University. His research interests include computer vision, face recognition and intelligent learning, in particular, focusing on person re-identification for visual surveillance. E-mail: chenypmail@163.com.

Xuekuan Wang is currently a master candidate in College of Electronics and Information Engineering, Tongji University. His research interests include computer vision, pattern recognition and machine learning, in particular, focusing on person re-identification for visual surveillance. E-mail:wxktongji@163.com.



Wai Keung Wong received his Ph.D. degree from The Hong Kong Polytechnic University. Currently, he is with Institute of Textiles & Clothing, The Hong Kong Polytechnic University, Hong Kong and The Hong Kong Polytechnic University Shenzhen Research Institute. He has published more than fifty scientific articles in refereed journals, including IEEE Transactions on Neural Networks and Learning Systems, Pattern Recognition, International Journal of Production Economics, European Journal of Operational Research, International Journal of Production Research, Computers in Industry, IEEE Transactions on Systems, Man, and Cybernetics, among others. His recent research interests include artificial intelligence, pattern recognition, and optimization of manufacturing scheduling, planning and control.

Duoqian Miao is currently a full professor and vice dean of the school of Electronics and Information Engineering of Tongji University. He received his PhD in Pattern Recognition and Intelligent System at Institute of Automation, Chinese Academy of Sciences in 1997. He works for Department of Computer Science and Technology of Tongji University, Computer and Information Technology Teaching Experiment Center, and the Key Laboratory of "Embedded System and Service Computing" Ministry of Education. He has published over 180 scientific articles in international journals, books, and conferences. He is committee member of International Rough Sets Society, senior member of China Computer Federation (CCF), committee member of CCF Artificial Intelligence and Pattern Recognition, committee member of Chinese Association for Artificial Intelligence (CAAI), chair of CAAI Rough Set and Soft Computing Society and committee member of CCAI Machine Learning, committee member of Chinese Association of Automation (CAA) Intelligent Automation, committee member and chair of Shanghai Computer Society (SCA) Computing Theory and Artificial Intelligence. His current research interests include: Rough Sets, Granular Computing, Principal Curve, Web Intelligence, and Data Mining etc.

Jingsheng Lei is a Professor and the Dean of the College of Computer Science and Technology, Shanghai University of Electronic Power. His research interests include web information retrieval, machine learning, data mining, and cloud computing. He received his B.S. in Mathematics from Shanxi Normal University in 1987, and M.S. and Ph.D. in Computer Science from Xinjiang University in 2000 and 2003 respectively.